DOCUMENT RESUME

ED 448 178 TM 032 139

AUTHOR Kennedy, Robert L.; McCallister, Corliss J.
TITLE Statistics Scores and Testing Time.
PUB DATE 2000-11-15
NOTE 8p.; Paper presented at the Annual Meeting of the Mid-South Educational Research Association (28th, Bowling Green, KY, November 15-17, 2000).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *College Students; Higher Education; *Scores; *Statistics; Test Length; Test Results; *Time; Time Management

ABSTRACT

The purpose of this study was to investigate the relationship between the scores students earned on their statistics final examinations and the number of minutes students required to complete the exams. In a previous study, K. Bridges (1985) extended the range of interest in this relationship from a single study to a course-based series, examining data generated from three multiple choice introductory psychology examinations. Little evidence was found to support either linear or nonlinear relationships between performance and order of finish or time required. In this study, the tests contained two parts: mechanical calculations of statistics and interpretations of printouts. Students were allowed unlimited time to complete the tests, and their times ranged from 69 to 305 minutes. Eight graduate-level basic statistics classes from fall terms between 1996 and 1999 were involved in the study, and their pooled group size, for which data were complete, was 97 (69 females and 28 males). The assumption of normality was not consistently met, so Spearman correlations were calculated: -0.30 for females ($p=0.01$), 0.24 for males ($p=0.22$), and -0.15 overall ($p=0.14$). The effect sizes for "r," according to Cohen, were medium for the females and small for the males and overall. That is, for the female students in this study, those who finished more quickly tended to perform better. For males, those who completed the tests quickly did not necessarily do better than those who took greater amounts of time, although there was a small tendency for those who took more time to perform better. Overall, the relationship was minimal. (Contains 16 references.) (Author/SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 448 178

Statistics Scores and Testing Time

Robert L. Kennedy
(rlkennedy@ualr.edu)

Corliss J. McCallister

Department of Educational Leadership

University of Arkansas, Little Rock

2801 S. University, Little Rock AR 72204-1099

Mid-South Educational Research Association

Twenty-Ninth Annual Meeting

University Plaza Hotel

Bowling Green, Kentucky

November 15, 2000

TM032139

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

R. Kennedy

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

BEST COPY AVAILABLE

Abstract

The purpose of this study was to investigate the relationship between the scores students earned on their statistics finals with the number of minutes the students required to complete the exams. In a review of the literature, Bridges noted a growing interest in the relationship between achievement test performance and order of finish as well as the time required, but observed that all of the studies he reviewed were based on single administrations of examinations. To extend the range of interest from a single test to a course-based series he examined data generated from three multiple-choice introductory psychology examinations. Little evidence was found to support either linear or nonlinear relationships between performance and order of finish or time required. In the present study, the tests comprised two parts: mechanical calculations of statistics and interpretations of printouts. The amount of time allowed the students to complete the tests was unlimited, ranging from 69 to 305 minutes. Eight graduate-level basic statistics classes from the fall terms, 1996-1999, were involved in the study and their pooled group size for which there was complete data was 97, including 69 females and 28 males. The assumption of normality was not consistently met, so Spearman correlations were calculated: -0.30 for females ($p=0.01$), 0.24 for males ($p=0.22$), and -0.15 overall ($p=0.14$). The effect sizes for r, according to Cohen, were medium for the females and small for the males and overall. That is, for the female students in this study, those who finished more quickly tended to perform better. For males, those who completed the tests quickly did not necessarily do better than those who took greater amounts of time, although there was a small tendency for those who took more time to perform better. Overall, the relationship was minimal.

Being the first person to finish a test has always appealed to some students. Whether it is competitive spirit, the ability to leave class early, or maybe to impress friends, there are always several students who will push themselves to be among the first to complete examinations. On the other hand, there are always a few students who tend to take whatever time is provided to complete tests. In the interest of insuring that an adequate amount of time is provided for students, it seems reasonable to question whether there is any relationship between the performance of students on tests and the amount of time required to complete the exams. In other words, do students really improve their scores when they are allowed extra time to complete examinations? The purpose of this study, then, was to investigate the relationship between the scores students earned on their final exams with the number of minutes the students required to complete the tests.

Several researchers have investigated performance and different aspects of testing, including the order of items, the location on the test of the response space, the handedness of the examinees, the sex of the examinees, and whether they belong to culturally-defined groups (Araujo & Semp, 1979; Donlon, 1977; Kleinke, 1979, 1980; Reilly & Evans, 1974). It was the observation of this researcher, however, that time may have some relationship with performance. In fact, it seemed logical that students who are well prepared would tend to be able to answer questions relatively quickly. In such a case, there may be no inherent advantage to allowing students extensive amounts of time to complete their tests.

In a study of 291 introductory general psychology students, Paul and Rosenkoetter (1980) noted a general trend for better students to finish examinations faster. They also found, however, that a disproportionate number of low-scoring students finished quickly and a similarly disproportionate number of high-scoring students were among the slower completers.

Wild, Durso, & Rubin (1982) compared scores on verbal and quantitative tests when students were given 20 minutes versus 30 minutes and found that there were no statistically significant differences for sex, race (black and white), or years out of school. This study closely replicated an earlier study (Wild and Durso, 1979) considering ethnicity, age, and sex in which only small improvements in performance were found with the increased time. The overall effect was less than a point.

Bridges (1982, 1985) did investigate order of finish and time required on tests. In his review of the literature, several studies were cited in which some significant curvilinear relationships had been found, but no linear relationships. He noted a growing interest in the relationship between achievement test performance and order of finish as well as the time required, but observed that all of these studies were based on single administrations of examinations. To extend the range of interest from a single test to a course-based series he examined data generated from three multiple-choice examinations given during five sections of introductory psychology offered by The Pennsylvania State University. Little evidence was found to support either linear or nonlinear relationships between performance and order of finish or time required.

Wright (1984) studied the effects of increased time limits on a college-level academic skills test. With 181 Miami-Dade Community College students participating, no significant differences were found for groups by native language, ethnicity, or sex. Gallagher and others (1985) studied possible bias in a similar Florida test, the College Level Academic Skills Test (CLAST), to see if non-native English speakers were at a disadvantage due to time constraints in taking the exam. The authors concluded that there was no significant relationship between extra time, ethnicity, and performance. Belcher and Einspruch (1987) allowed subjects twice as long, 70 minutes as opposed to 35 minutes, to retake CLAST subtests, and found that they outperformed the subjects having the lesser amount of time.

Grover and others (1988) studied 50 college students enrolled in introductory psychology in terms of test-taking time, performance, and personality behavior type. Their findings indicated that test-taking time is not related to personality type.

Hale (1992) examined essay performance on the Test of Written English by students with 30 minutes versus 45 minutes for completion. A total of 820 intensive English and academic international students participated. Mean increases for the additional 15 minutes were less than half a point.

Lawrence (1993) focused on Scholastic Aptitude (SAT) and Graduate Management Admissions Tests and considered the effect of test speededness for race and sex. Subgroup differences in test scores were found to be minimal, although it did appear that black and Hispanic examinees received credit for more difficult sets of items in the SAT verbal test, relative to whites.

In the present study, the tests comprised two parts: mechanical calculations of statistics and interpretations of printouts. Tests were administered at the end of each fall term beginning with 1996 and continuing through 1999. Students were allowed whatever time they desired to make their selections for the twenty items. When their tests were turned in, the time was noted and recorded on the test. When the tests were graded, the scores were the total number of questions answered correctly, with no weighting.

The amount of time taken by the students to complete the tests ranged from 69 to 305 minutes. Eight graduate-level basic statistics classes from the 1996-1999 fall terms were involved in the study and their pooled group size for which there was complete data was 97, including 69 females and 28 males. The assumption of normality was not consistently met, so Spearman correlations were calculated: -0.30 for females ($p=0.01$), 0.24 for males ($p=0.22$), and -0.15 overall ($p=0.14$). The effect sizes for r, according to Cohen, were medium for the females and small for the males and overall. That is, for the female students in this study, those who finished more quickly tended to perform better. For males, those who completed the tests quickly did not necessarily do better than those who took greater amounts of time, although there was a small tendency for those who took more time to perform better. Overall, the relationship was minimal, consistent with most of the findings noted in the literature review.

References

Araujo, J., & Semb, G. (1979, September). The Effects of Item Order on Student Performance. Paper presented at the Annual Meeting of the American Psychological Association, New York, NY. (ERIC Document Reproduction Service No. ED 185 086)

Belcher, M. J., & Einspruch, E. (1987). An examination of CLAST retake performance: Is more time an advantage to test takers? (ERIC Document Reproduction Service No. ED 323 237)

Bridges, K. R. (1982, April). Order of Finish, Time to Completion, and Performance on Course-Based Objective Examinations: A Systematic Analysis of Their Relationship on Both Individual Exam Scores and Total Score Across Three Exams. Paper presented at the Annual Meeting of the Eastern Psychological Association, Baltimore, MD. (ERIC Document Reproduction Service No. ED 223 661)

Bridges, K. R. (1985, Spring). Test-Completion Speed: Its Relationship to Performance on Three Course-Based Objective Examinations. Educational and Psychological Measurement, 45,(1), 29-35.

Donlon, T. F. (1977, October). Sex Differences in Test Speededness on the Scholastic Aptitude Test. Paper presented at the Annual Meeting of the Northeastern Educational Research Association, Ellenville, NY. (ERIC Document Reproduction Service No. ED 163 047)

Gallagher, P., and others. (1985). Time as a factor in test performance in a state mandated testing program. (ERIC Document Reproduction Service No. ED 275 729)

Grover, C. A., and others. (1988, April). Test-taking time, performance, and the Type A behavior pattern. (ERIC Document Reproduction Service No. ED 297 223)

Hale, G. A. (1992, June). Effects of amount of time allowed on the test of written English. (ERIC Document Reproduction Service No. ED 385 569)

Kleinke, D. J. (1979, April). Item Order, Response Format, and Examinee Sex and Handedness and Performance on a Multiple-Choice Test. Paper presented

at the Annual Meeting of the National Council on Measurement in Education, San Francisco, CA. (ERIC Document Reproduction Service No. ED 171 793)

Kleinke, D. J. (1980, March-April). Item Order, Response Location, and Examinee Sex and Handedness and Performance on a Multiple-Choice Test. Journal of Educational Research, 73,(4), 225-29.

Lawrence, I. M. (1993, October). The effect of test speededness on subgroup performance. (ERIC Document Reproduction Service No. ED 171 793)

Paul, C., & Rosenkoetter, J. (1980). Relation between completion time and test score. Southern Journal of Educational Research, 14(2), 151-57.

Reilly, R. R., & Evans, F. R. (1974, September). The Effects of Test Time Limits on Performance of Culturally Defined Groups. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA. (ERIC Document Reproduction Service No. ED 102 276)

Wild, C., & Durso, R. (1979, June). Effect of increased test-taking time on test scores by ethnic groups, age, and sex. (ERIC Document Reproduction Service No. ED 241 570)

Wild, C. L., Durso, R., & Rubin, D. B. (1982, Spring). Effect of increased test-taking time on test scores by ethnic group, years out of school, and sex. Journal of Educational Measurement, 19(1), 19-28.

Wright, T. (1984, April). The effects of increased time limits on a college-level achievement test. Research report no. 84-12. (ERIC Document Reproduction Service No. ED 267 867)